Production of Demonstratives in Dutch, English and Portuguese dialogues

van der Sluis, Ielka; Luz, S.

Published in:
Proceedings of the 13th European Workshop on Natural Language Generation (ENLG'11)

Publication date:
2011

Production of Demonstratives in Dutch, English and Portuguese Dialogues

Saturnino Luz
Department of Computer Science
Trinity College Dublin, Ireland
luzs@cs.tcd.ie

Ielka van der Sluis
Communication and Information Studies
University of Groningen, the Netherlands
I.F.van.der.Sluis@rug.nl

Abstract

A data elicitation study on the type of demonstratives and determiners selected to denote objects in English, Dutch and Portuguese dialogues is presented. Participants were given a scenario and a scripted dialogue in which a furniture seller identifies target objects to a buyer. They were then asked to choose a combination of a determiner or demonstrative and a referring expression to be uttered by the seller and told that the agent would point at the targets while uttering the chosen linguistic descriptions. The study was conducted with native speakers and rendered a total of 920 demonstratives and determiners. It focused on accessibility of the target referents and distance between agents and target referents. Results show that the three language groups largely agree in their preferences and, in contrast to previous work, align with a nearby/far away distinction.

1 Introduction

This paper investigates the use of indexical determiners (i.e. determiners employed for direct references to objects and that include a pointing gesture) by Dutch, Portuguese and English speakers. A comparison of the use of Dutch and English demonstratives in terms of the accessibility of the target by Piwek and Cremers (1996) suggested that English and Dutch speakers follow opposite strategies in their use of indexical demonstratives. Dutch speakers use proximal demonstratives for referents that are relatively difficult to access (*deze*), while English speakers use proximals (*this*, *these*) for referents that are relatively easy to access. Piwek et al.(2008) present an explanation for these differences in terms of the use of pointing gestures (Clark and Bangerter, 2004; Bangerter, 2004), suggesting that a pointing gesture functions as a "labelling" of the target object as being relatively accessible. Hence, where proximals require a pointing gesture, distal demonstratives (*dat*/*that* and *die*/*those*, which are more similar to definite determiners) can also be used non-indexically. This model corresponds to the 'folk-view' of demonstratives that considers distals to indicate objects far away from the speaker and proximals to indicate objects near the speaker (Bühler, 1934; Clark, 1996).

Byron and Stoia (2005) present a motivation for choosing either a proximal or a distal demonstrative based on three dimensions (i.e. spatial, temporal and task performance). Their analysis of a corpus of collaborative dialogues between participants solving a treasure hunt problem in a virtual space, shows that, in English: (1) distals are used both for objects located close to and far away from the speaker, whereas proximals are only used for objects located near the speaker; (2) proximals are used for objects that relate to the current time and to the future, while distals are used for past events; and (3) distals are less sensitive to the space and time dimension and more sensitive to the task than proximals.

While we acknowledge that these are important dimensions in the analysis of demonstratives, in the present paper we restrict ourselves to an elicitation study and analyse the use of indexical determiners in terms of accessibility and distance, in line with the model developed by Piwek et al. (1996; 2008).

Proceedings of the 13th European Workshop on Natural Language Generation (ENLG), pages 181–186,
Nancy, France, September 2011.

In addition to the languages compared by Piwek et al. (Dutch and English), we also analyse the use of demonstratives in Portuguese. The dialogue context designed for this study fits a discursive context in which the distal/proximal distinction is appropriate for Portuguese (Cavalcante, 2002), and thus enables us to compare the use of demonstratives across these three languages. Another difference between this study and those conducted by Piwek et al. is the data gathering method. While those authors relied on corpora collected from free task-based dialogues between participants, we employed scripted dialogues (André et al., 2000; Williams et al., 2007) presented to individual participants who were explicitly asked to choose among demonstratives.

2 Production Study

The study presented below originated from an investigation into the perception of multimodal referring expressions (REs) in a virtual world by Japanese an English speakers (Van der Sluis and Luz, 2011b; Van der Sluis et al., to appear). In this paper, the materials from a production study initially conducted for Japanese to validate our Japanese translation of a dialogue written in English, have been translated and further adapted to Dutch and Brazilian Portuguese. We draw on the results of this study to analyse the use of demonstratives in English, Dutch and Portuguese. The REs considered in this study are embedded in a scripted dialogue between two agents in a furniture sales setting. The study focuses on 'first-mention' REs which identify objects that have not been talked about earlier in the discourse.

A dialogue script was written for two agents in a furniture shop. The layout of the shop and the positions of the agents and furniture items is shown in Figure 1. The shop contains 26 objects, comprising distractors as well as target referents. The dialogue consists of 19 utterances and features a conversation between a female agent purchasing furniture for her office, and a male shop-owner describing some furniture items. The furniture seller agent refers to objects in the domain by uttering each scripted RE combined with a pointing gesture directed to the target object. Validation showed that the dialogue was acceptable to English speakers. Van der Sluis and Luz (2011a) describe the setting in greater detail.

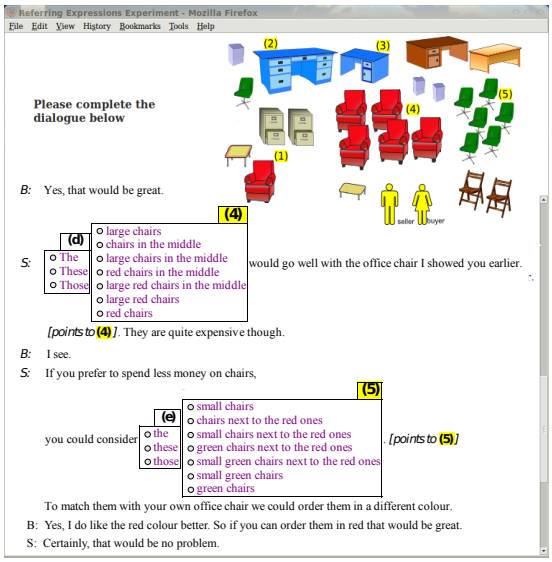

Figure 1: Screenshot of the application in which participants were asked to choose their preferred REs. Utterances by the Seller and Buyer are marked with "S:" and "B:", respectively. Options were presented as shown in the DE-boxes marked *(d)* and *(e)*, and RE-boxes marked *(4)* and *(5)*.

The dialogue was used as a template in which five first-mention referring expressions (REs) could be varied. The REs used to fill out these slots were chosen to cover various aspects of REs as are currently being studied in NLG: (1) cardinality, the REs targeted three singular objects and two larger sets of items; (2) locative expressions, the REs included three absolute locative expressions and two relative locative expressions; and (3) the position of the referent, the targets were distributed in the domain of conversation such that one referent was located near to the stationary agents, two referents were located far away from the agents, and two sets of referents were located somewhere in between those two extremes. Figure 1 shows 14 furniture items that are used for assessing multimodal GRE output: labelled (1) to (5), as well as a number of distractors. It was assumed that the agents would stay stationary and point in the direction of the targets.

The text was translated into Dutch and Brazilian Portuguese so as to adapt the dialogue to the normative, communicative and inferential rules of the respective cultures, but we attempted to keep the REs as close to the English originals as possible. The translations and localisations for Dutch and

Portuguese followed a similar pattern as the process for Japanese described in (Van der Sluis and Luz, 2011b). Validation of the translated dialogues was conducted by three native speakers in the respective languages and revisions were made accordingly.

Although the study was also conducted for Japanese, we will restrict our discussion in this paper to Dutch, English and Portuguese because the Japanese system for demonstratives differs from the ones discussed in this paper. It is a ternary, person-oriented system (Anderson and Keenan, 1985, p.282-286), in contrast to distance-oriented system such as the ones that seem to govern the use of demonstratives in Portuguese, English and Dutch. Although the Portuguese system also incorporates three classes of demonstratives, namely: *este(a)/s*, *isto*, (proximal), *esse(a)*, *isso (medial)* and *aquele(a)/s* (distal), these often operate as a binary system where the *este* and *esse* classes are used interchangeably as proximals whereas *aquele* is used as a distal (Cavalcante, 2002; Jungbluth, 2005).

Linguistic preferences were elicited through a web-based application. After being introduced to the scenario and task, participants were shown a screen similar to Figure 1. A picture of the domain was displayed at the top and kept visible throughout the dialogue. The bottom part of the screen contained the dialogue, through which the participants could scroll and select the REs and determiners or demonstratives from a set of options, all of which were simultaneously available to the participant while reading the sentence. The five REs were each presented with two boxes as illustrated in Figure 1: the DE-box, for determiner or demonstrative selection and the RE-box, for referring expressions. After each RE-box, it was stated that the agent's utterance would be combined with a pointing gesture in the direction of the target. The REs collected with the study are analysed elsewhere (Van der Sluis and Luz, 2011a) and will not be further considered in this paper. The DE-box included three options for Dutch, English and Portuguese: a definite determiner and a proximal and distal demonstrative.

3 Hypotheses

Two hypotheses, denoted H1 and H2 and summarised in Table 1, were tested for the five REs produced in the dialogue with respect to the use indexical demonstratives. H1 is related to the accessibility of the target (Gundel et al., 1993) and H2 concerns the physical distance between the speaker and the target object. Compared to the targets of RE2 and RE4, objects identified by RE1, RE3 and RE5 are relatively easier to access because they are located in the 'focus area' of the discourse (RE3 and RE5) or set visibly apart from the other objects in the domain (RE1). Hence, RE1, RE3 and RE5 call for demonstratives that indicate easy access. According to Piwek and Cremers (1996), Dutch speakers prefer proximal demonstratives for objects which are relatively hard to access, while English speakers apparently follow the opposite strategy. Portuguese speakers appear to follow a strategy which is similar to the latter (Cavalcante, 2002). In order to test these claims we set the accessibility hypothesis, H1, so that it predicts opposite strategies for Dutch, on the one hand, and Portuguese and English on the other.

Table 1: Expected *prox*imal and *dist*al demonstratives for *E*nglish and *D*utch for REs 1 to 5 with respect to ease of *a*ccess, (*H1*) and *d*istance, (*H2*).

RE	*a*	*H1-E/P*	*H1-D*	*d*	*H2-EDP*
RE1	easy	prox	dist	near	prox
RE2	difficult	dist	prox	far	dist
RE3	easy	prox	dist	far	dist
RE4	difficult	dist	prox	near	prox
RE5	easy	prox	dist	far	dist

Hypothesis H2 relates to the distance between target object and speaker. It predicts that participants will prefer distals over proximals to indicate objects further away (i.e. a proximal for RE1 and RE4 and distal demonstratives for the other REs). Since the dialogue script includes an explicit pointing gesture for all REs, we expected participants to choose either a proximal or an (indexical) distal demonstrative. We had no hypotheses about the use of definite determiners and exclude them from further analysis.

4 Results

Participants included 91 native English speakers (60% female, 40% male; age groups: 52% between 20 and 30, 33% between 31 and 40, and 25% over 41 years old; occupations: 44% students, 26% academics and 31% other), 42 native Brazilian Portuguese speakers (female: 60% female, 40% male;

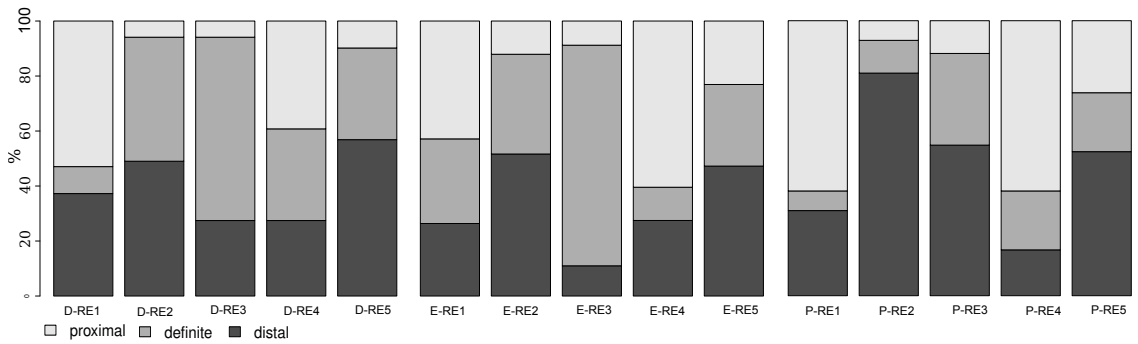

Figure 2: Percentages of definite determiners, distal and proximal demonstratives per referring expression (RE1 to RE5) for Dutch, English, and Brazilian Portuguese.

age groups: 71% between 20 and 30, 26% between 31 and 40, and 2% over 41 years old; occupations: 29% students, 57% academics and 14% other) and 51 native Dutch speakers (female: 55% female, 45% male; age groups: 21% between 20 and 30, 33% between 31 and 40, and 26% between 41 and 50 and 20% over 50 years old; occupations: 4% students, 14% academics and 82% other).

4.1 Demonstratives

Figure 2 presents the percentages of definite determiners, proximal and distal demonstratives selected per RE per language. Results show that native speakers of Portuguese, Dutch and English roughly agree in their choices. However, for RE3 we found some disagreement. The majorities of the English and Dutch participants, did not select a demonstrative, but selected a definite determiner for RE3 (i.e. 'the small desk next to it'). In contrast, the Portuguese speakers preferred a distal demonstrative.

Table 2 shows the frequencies of the demonstratives selected, determiners excluded. Again, Portuguese, Dutch and English speakers mostly agree in their choices. The majorities chose a proximal demonstrative for RE1 (i.e. 'this red chair'), a distal demonstrative for RE2 (i.e. 'that large desk'), a distal for RE3 (i.e. 'that small desk next to it'), a proximal demonstrative for RE4 (i.e. 'these red chairs') and a distal demonstrative for RE5 (i.e. 'those green chairs next to the red ones').

We computed χ^2 statistics to assess whether the data borne out the differences hypothesised (Table 1) and if those differences were statistically significant (i.e. whether the null hypotheses that no difference exists could be confidently rejected). The results of these tests are also summarised in Table 2.

Table 2: Frequencies of definite *Det*erminers and *Distal* and *Proximal* demonstratives per *RE*ferring expressions for the *L*anguages *E*nglish, *D*utch and Brazilian *P*ortuguese, where differences are indicated with * = $p < .05$ and ** = $p < 0.01$. Where the null hypothesis is rejected, a + sign indicates a difference that agrees with the alternative hypothesis (H1, H2), and a − sign indicates a difference that disagrees with the alternative hypothesis.

RE	*Distal*	*Proximal*	*H1*	*H2*
E-RE1	38%(24)	62%(39)		
E-RE2	81%(47)	19%(11)	+**	+**
E-RE3	56%(10)	44%(8)		
E-RE4	31%(25)	69%(55)	−**	+**
E-RE5	67%(43)	33%(21)	−**	+**
D-RE1	41%(19)	59%(27)		
D-RE2	89%(25)	11%(3)	−**	+**
D-RE3	82%(14)	18%(3)	+**	+**
D-RE4	41%(14)	59%(20)		
D-RE5	85%(29)	15%(5)	+**	+**
P-RE1	33%(13)	67%(26)	+*	+*
P-RE2	92%(34)	8%(3)	+**	+**
P-RE3	82%(23)	18%(5)	−**	+**
P-RE4	21%(7)	79%(26)	−**	+**
P-RE5	67%(22)	33%(11)		

English participants agreed with our Access hypothesis and Distance hypothesis for RE2 ($\chi^2[1] = 22.35, p < .01$), which predict a distal demonstrative. English participants agreed with the Distance hypotheses for RE4 ($\chi^2[1] = 11.25, p < .01$) and RE5 ($\chi^2[1] = 7.56, p < .01$) and rejected the Access hypotheses for these REs (i.e. respectively a

proximal and a distal demonstrative were preferred for RE4 and RE5). Dutch participants chose distal demonstratives for RE2, RE3 and RE5, respectively ($\chi^2[1] = 17.29, p < .01$), ($\chi^2[1] = 7.12, p < .00$) and ($\chi^2[1] = 16.94, p < .01$), thereby agreeing with the Distance and also with the Access hypothesis for RE3 and RE5. However, for RE2 the Dutch participants disagreed with the Access hypothesis. Portuguese speakers agreed with the Access hypothesis and the Distance hypothesis for RE1 ($\chi^2[1] = 4.333, p < .05$) and RE2 ($\chi^2[1] = 25.97, p < .01$) preferring respectively a proximal and a distal demonstrative. They also agreed with the Distance hypothesis for RE3 ($\chi^2[1] = 11.57, p < .01$) and RE4 ($\chi^2[1] = 10.94, p < .01$), preferring respectively a distal and proximal demonstrative, and thus rejected the Access hypothesis.

4.2 Access versus Distance

Table 3 summarises the participants' choices in terms of Access (H1) and Distance (H2) for the three language groups in the cases where the hypotheses differed. For English participants such differences were found for RE4 ($\chi^2[1] = 11.25, p < .01$) and RE5 ($\chi^2[1] = 7.56, p < .01$) indicating that their selections matched the Distance hypothesis better than the Access hypothesis. The Dutch participants also matched the Distance hypothesis better but only for RE2 ($\chi^2[1] = 17.29, p < .01$). Finally the demonstratives selected by the Portuguese participants matched the Distance hypothesis for RE3 ($\chi^2[1] = 11.57, p < .01$) and RE4 ($\chi^2[1] = 10.94, p < .01$) better than the Access hypothesis.

Table 3: Successful predictions of demonstratives for hypotheses *H1* (accessibility) and *H2* (distance) for *E*nglish, Brazilian *P*ortuguese and *D*utch. Significant differences between H1 and H2 are denoted with '**' ($p < .01$).

RE	*H1-Access*	*H2-Distance*	*H1 vs H2*
E-RE3	44%(8)	56%(10)	
E-RE4	31%(25)	69%(55)	**
E-RE5	33%(21)	67%(43)	**
D-RE1	41%(19)	59%(27)	
D-RE2	11%(3)	89%(25)	**
P-RE3	18%(5)	82%(23)	**
P-RE4	21%(7)	79%(26)	**
P-RE5	33%(11)	67%(22)	

5 Discussion and Conclusion

The Distance Hypothesis (H2) appears to be a better fit to the preferences of native speakers of the three languages than the Accessibility Hypothesis (H1). It agrees with the majority of choices for RE1, RE2, RE4 and RE5 in all language groups. Expression RE3, however, proved to be something of an exception, specially for Dutch and English, in that participants of those languages preferred to use a definite determiner in this RE rather than a distal or proximal demonstrative. It seems that in this case the increased accessibility of object (3), caused by the previous reference to 'the desk next to it', was transferred to the definite determiner rather than the distal demonstrative for Dutch and the proximal demonstrative for English, as predicted by H1.

In contrast to previous work, the data collected in our study show that the majorities of the three languages agree in their choices of demonstratives. This may be explained by the fact that pointing gestures were an explicit part of the REs that we tested, and therefore could be evidence for the post-hoc analysis presented by (Piwek et al., 2008), aligning with the folk view of a nearby/far away distinction.

Finally, this study introduced some methodological innovations. Unlike studies where data are collected from naturalistic conversations, we explicitly asked participants to make a judgement as to which demonstrative to use. This was done in a context which, although arguably still open to subjective interpretation, is much more tightly controlled and therefore better suited to cross-linguistic comparison. However, it could be argued that better control comes at the cost of naturalness. By asking the participants to respond from a third person's perspective and imagine the effects of communicative acts (including gestures) the study might have favoured reflective answers over spontaneous production. Such trade-offs seem to be characteristic of this sort of study, and getting them right is one of the many challenges in language generation research.

Acknowledgements

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at Trinity College Dublin.

References

S. Anderson and E. Keenan. 1985. Deixis. In *Language Typology and Syntactic Fieldwork Vol. III*. Cambridge: Cambridge University Press.

E. André, T. Rist, S. Mulken, and M. Van Klesen. 2000. The automated design of believable dialogues for animated presentation teams. In *Embodied Conversational Agents*. MIT Press.

A. Bangerter. 2004. Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, 15:415–419.

K. Bühler. 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache.* Fischer, Jena.

D. Byron and L. Stoia. 2005. An analysis of proximity markers in collaborative dialog. In *Proceedings of CLS-05*.

M. Cavalcante. 2002. O demonstrativo e seus usos. *Perspectiva*, 20(1).

H. Clark and A. Bangerter. 2004. Changing ideas about reference. In I. Noveck and D. Sperber, editors, *Experimental Pragmatics*, pages 25–49. Palgrave Macmillan, New York.

H. Clark. 1996. *Using Language*. Cambridge University Press.

J.K. Gundel, N. Hedberg, and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.

K. Jungbluth. 2005. Os pronomes demonstrativos do português brasileiro na fala e na escrita. *Cadernos de Linguagem e Sociedade*, 7:83–105.

P. Piwek and A. Cremers. 1996. Dutch and English demonstratives: A comparison. *Language Sciences*, 18(3-4):835–851.

P. Piwek, R. Beun, and A. Cremers. 2008. 'Proximal' and 'distal' in language and cognition: Evidence from deictic demonstratives in Dutch. *Journal of Pragmatics*, 40:694–718.

I. Van der Sluis and S. Luz. 2011a. A cross-linguistic study on the production of multimodal referring expressions in dialogue. In *Proceedings of the 13th European Workshop on Natural Language Generation*, Nancy, France.

I. Van der Sluis and S. Luz. 2011b. Issues in translating and producing Japanese referring expressions for dialogues. *Linguistic Issues in Language Technology*, 5(1):1–46.

I. Van der Sluis, S. Luz W. Breitfuß, M. Ishizuka, and H. Prendinger. to appear. Cross-cultural assessment of automatically generated multimodal referring expressions in a virtual world. *International Journal of Human-Computer Studies*.

S. Williams, P. Piwek, and R. Power. 2007. Generating monologue and dialogue to present personalised medical information to patients. In *Proc. of ENLG-07*.